HEP computing at ANL

S.Chekanov (ANL)

HEP DOE Scientific Computing program review at ANL

Feb 8-10

Plan of this talk

- **Introduction to the HEP division (see Harry's talk)**
- **Budget model**
- **Division's computing resources**
- **Scientific HEP computing:**
  - General computer-infrastructure support
  - Experiment-specific computing infrastructure support
  - Experiment-specific software development
  - Software solutions for general scientific HEP computation
- **Future**

HEP ANL

- **Astrophysics**
- **Accelerator physics**
- **Theory**
- **Electronic/Mechanical support**
- **Detector R&D**
- **CDF group**
- **Neutrino**
- **ATLAS** group and **ATLAS Analysis Support Center (ASC)**

  to provide US ATLAS physicists with regional resources, tutorials, support, leadership and focal point for meetings
- **ATLAS** computing group (See D.Malon's talk)
  - Event store, I/O framework, persistence technologies, distributed database services, metadata etc.

~60 ANL/HEP physicists
~40 ASC users (25 non-ANL)

Division's role:

- detector construction, detector R&D for HEP experiments
- data analysis & physics simulation (ATLAS, Neutrino, Astro groups)
- software development

HEP ANL budget model

- **Use DOE HEP's research and technology R&D ("base" funds) for:**
  - **General computing infrastructure**
    - ~2 FTE to provide computing services
    - general hardware & software support:
      - $55k (60% paid for MS licenses, the rest -hardware for replacement)
  - **Experiment specific infrastructure support**
    - Typically $10-20k from each group
    - *Recently: ARRA money for ATLAS Tier3 equipment:*
      - *Last year: $214k one-time supplemental for ATLAS computing (Tier3 + ASC)*
      - *This year: $40k for ATLAS 10Gb network installation in Bld. 360*
- **US ATLAS Operations Program to support 5 FTEs** (see D.Malon's talk)
- **Argonne's Laboratory Directed Research and Development (LDRD) program**
  - experiment specific, but irregular

- The division does not host any major HEP facility (no HEP facility operations funding)
- No dedicated computer infrastructure to provide services for HEP experiments or computing service to HEP community (SPIRES, PDG etc.)

ANL HEP computing environment

Give perspectives from:

- "analysis-type" HEP facility (small HEP labs, Universities)
- Resource-limited:
  - no dedicated professional setup, professional computer room
  - often lacks trained computer scientists with HEP background

Main issues: How to

- Achieve division's goals having limited local resources & operations funds
- Leverage our resources using the Lab computing services
- Take advantage of the existing ANL computer & network infrastructure
  - requires significant intellectual investments to adopt for concrete HEP tasks

ANL HEP computer-group manpower

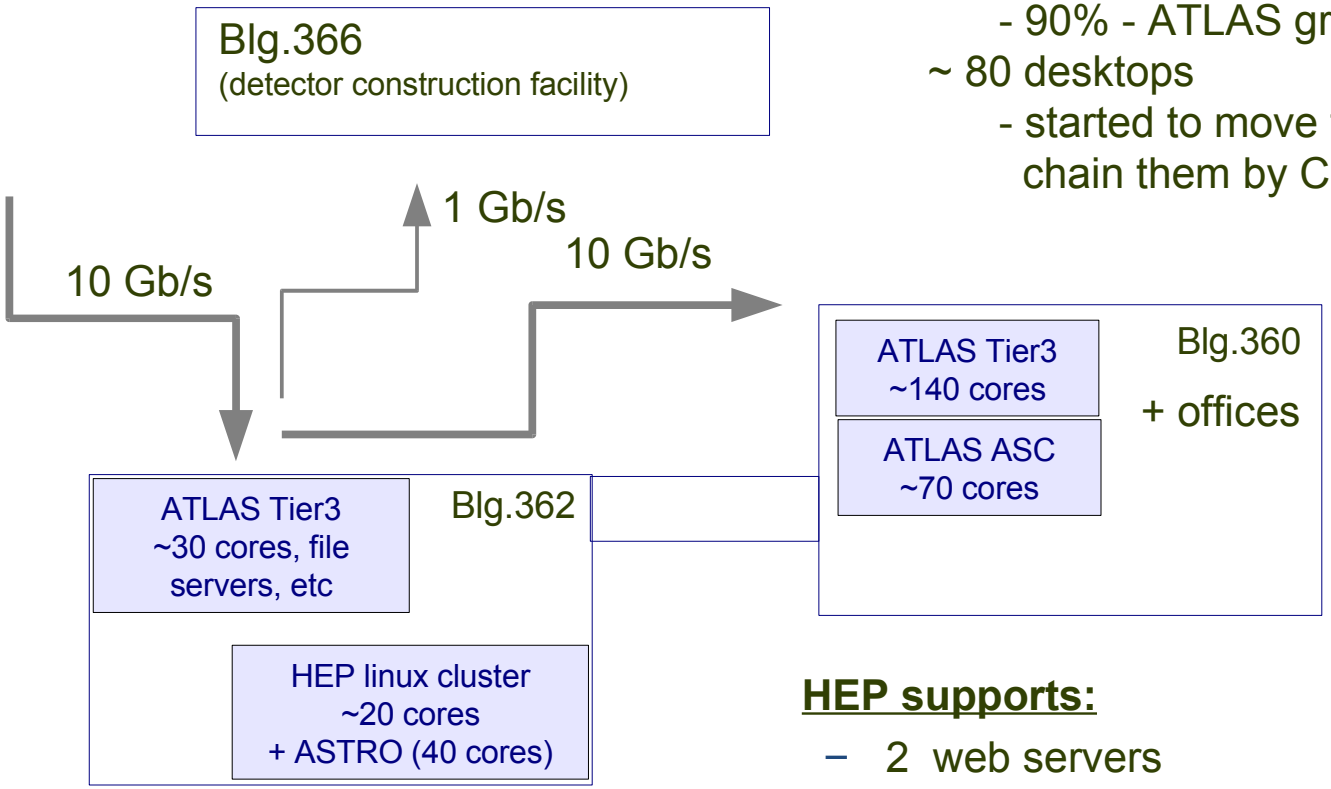
- **General computer support**
  - ~2 FTEs including 2 physicists (50-20% FTE each)
- **Experiment specific computer support:**
  - few physicists with 20% FTE each
  - R. Yoshida: US ATLAS Tier3 coordinator (will stop this year)
    + D.Benjamin (ANL visitor from Duke Univ.)
    ATLAS & USATLAS Tier3 technical coordinator
- **5 FTE - ATLAS computing group (USATLAS operations funds)**
  - Event store, I/O framework, persistence technologies, distributed database services, metadata, event selection services (→ next talk)
- **2 new FTE**
  - to bring computational cosmology program (→ next talk)

Division's computing resources

Total hardware:

~ 40 servers (about 320 CPUs)
- 90% - ATLAS group

~ 80 desktops
- started to move to i7 desktops and chain them by Condor



HEP supports:

- 2 web servers
  - *hep.anl.gov* (for all HEP)
  - *atlaswww.hep.anl.gov* (for ATLAS)
- mail server
- Indico, Twiki, a Document server

General computing infrastructure support

Main goals:
→ **support for all computing needs on daily bases a trouble-ticketing system**

- Windows & ANL-domain accounts
- Access to group-owned licenses (but paid by HEP)
- Printer support
- Supporting 3 Linux clusters (2 are ATLAS specific)
- Responsible for computer security, deploying software patches, protecting against security problems with preventive measures (shared with the Lab's CIS)
- Maintain general resources: Twiki (2), Indico, 2 web servers, mail servers
- Researches the suitability of applications for use by the physicist community
  - many installed packages are redesigned to meet our needs

+ significant help from CIS

General computing infrastructure support

- **CIS =** ~~Citizenship and Immigration~~ Services

- **CIS → Computing and Instrumentation Solutions Division (CIS)**
  - All cybersecurity issues at ANL:
    - cybersecurity monitoring, combat unsolicited e-mails, etc.
    - cybersecurity trainings:
  - Support for HEP Windows domain accounts
  - Network uplink, IP assignments, installation of network switches (but paid by us)
  - Started to use ANL authentication (via kerberos) for ATLAS/ASC Linux cluster

→ **HEP administrators are in direct contact with CIS (or via CIS HelpDesk)**

Experiment-specific computing infrastructure support

ATLAS group:

Involved in operation and maintenance of two Linux clusters for ATLAS:

- **ANL Tier3 cluster (with ~150 cores)**.
  - in operation since 2010
  - used by ATLAS group + ASC users affiliated at HEP ANL
- **ASC/integration cluster (~70 cores)**
  - full-scale proof-of concept Tier3(g) cluster
  - endorsed as a reference Tier3 model for data analysis
  - used to develop Tier3 documentation (now moved to CERN)
  - used by ASC users. In operation since 2010
  - constructed by
    - R.Yoshida (US ATLAS Tier3 coordinator) +S.C. (ANL Tier3)
      D.Benjamin (ATLAS Tier3 technical coordinator) - ANL visitor from Duke Univ.

ANL-HEP Tier3 ATLAS cluster (Bld. 360)

- 150 cores. 2.25 GB/core
- 75 TB local storage (20TB used already)

- "interactive" & "chaotic" analysis
- scalable, no I/O bottleneck:
  - processes tens of TBs data with almost no load on network
- running Condor and ArCond (**Ar**gonne **Cond**or) for submission of ATLAS packages

Ganglia monitoring

- Put together ~0.5FTE (~2 months)
- This year: 10 Gb fiber connection
- Current support: ~0.2 FTE

Cluster is split between 2 buildings (360 and 362)

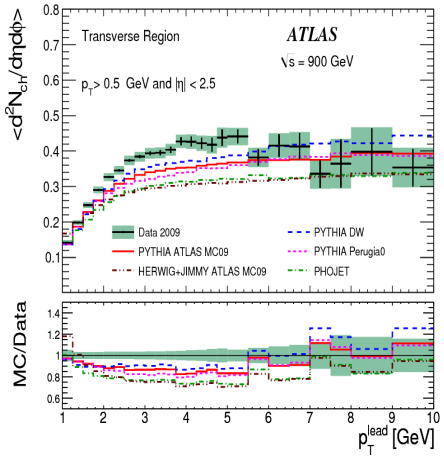
- several interactive nodes, 26 TB file server on 10 Gb network, NFS servers, web servers etc. are located in two labs of Bld. 362

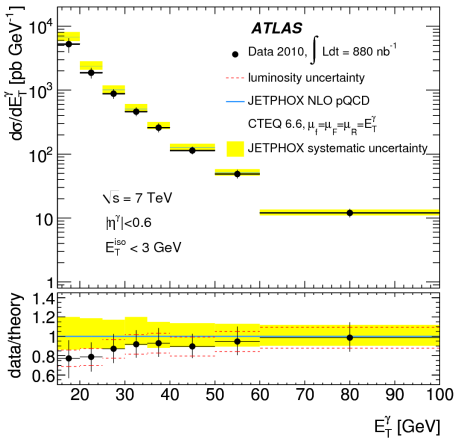
First physics results using the ANL Tier3 cluster

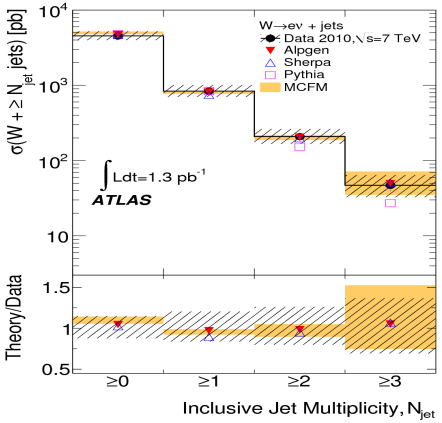
ATLAS papers:

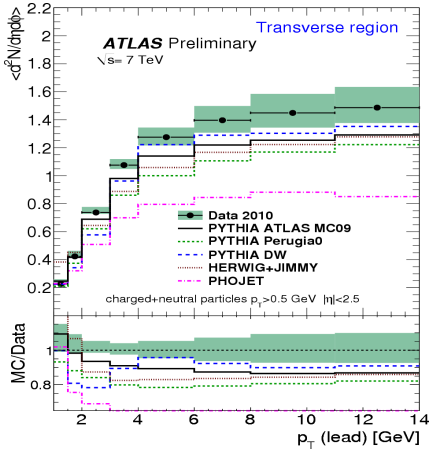
+ ~15 ATLAS notes

- **W+jet measurement** (arXiv:1012.5382)
  - ATLAS paper is out. ~50% done here
- **Underlying event analysis using tracks** (arXiv:1012.0791)
  - ATLAS paper is out. Complete second analysis
- **Inclusive direct photon production** (arXiv:1012.4389)
  - ATLAS paper is out. Theoretical calculations and scale uncertainties
- **Underlying event analysis using topological clusters**
  - ATLAS paper is about to be released. 100% done at ANL









Tier3 US ATLAS integration cluster (Bld. 360)

~70 cores 20 TB: ArCond, XROOTD, PROOF, gridFTP etc..

A low-cost well-balanced HEP cluster (tens $k) designed for US ATLAS

- "interactive" & "chaotic" analysis
- heavy I/O (processing tens of TB data)
- scalable, no I/O bottleneck

- complete documentations developed here and migrated to CERN

The cluster has been adopted as "example" Tier3(g) by US ATLAS as well as the wider ATLAS collaboration

- **R. Yoshida**
  - US ATLAS Tier3 coordinators. ASC integration cluster
- **D.Benjamin** (ANL visitor from Duke Univ.)
  - ATLAS Tier3 technical coordinator. ASC integration & test cluster

~44 ATLAS Tier3

+ HEP computer-group support

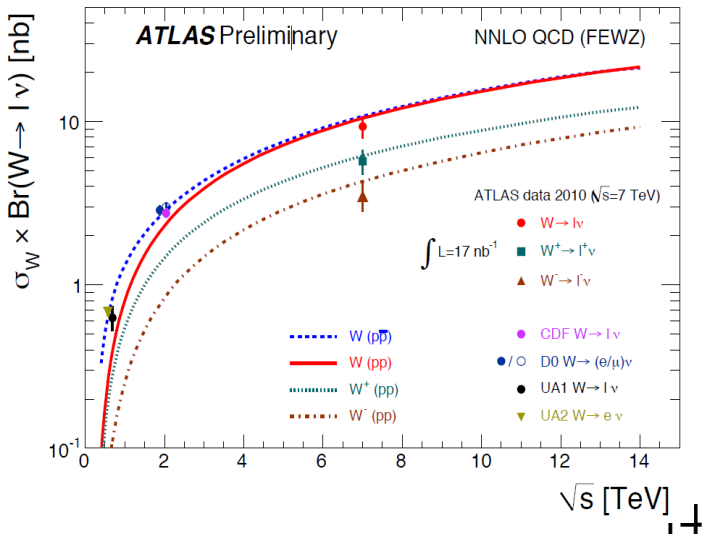
HEP computation

Theory group:

- **focuses on exploring strongly-interacting gauge theories, in particular "walking" technicolor** (D.Sinclair)
  - Total allocation ~8.5M CPU-hours (2010)
    - 4.2M CPU-hours on NERSC cluster (JERCAP DOE project)
    - 4 M CPU-hours on NICS (via TRAC-NSF)
    - currently low-level parallelism (no use of BlueGene/P, max ~300 cores)
    - Expectation for the next year: 10 M CPU-hours

- **Authored and maintains several analysis codes that are standard tools in the collider physics community such as:**
  - FEHiP (Fully Exclusive Higgs Production)
  - FEWZ (Fully Exclusive W and Z production)
  - Current requirement ~2 K CPU-hours
    - can be done using HEP resources
    - need to explore ANL Fusion cluster

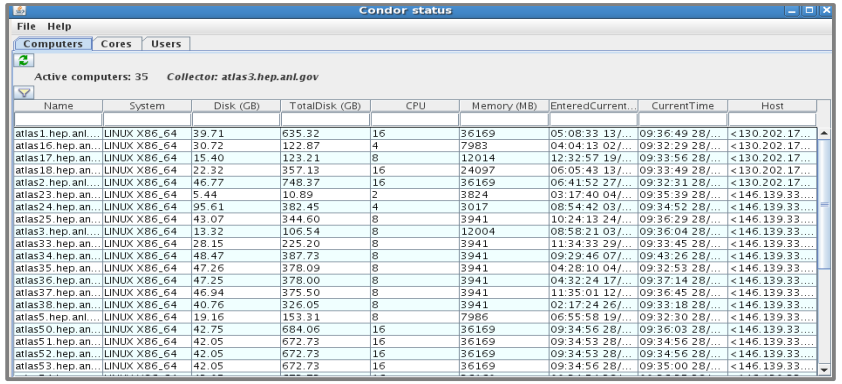
(F.Petriello, R.Boughezal,S.Quackenbush)



Experiment-specific software development

- **Neutrino group (MINOS (1987-2011), NOvA (2004-2011) and LBNE (2009-2011):**
  - Contributions to Monte Carlo simulations
  - Currently: Monte Carlo simulation for NovA (~75% FTE)
- **ATLAS group:**
  - **ArCond** (http://atlaswww.hep.anl.gov/asc/arcond/)
    - Condor front-end for I/O intensive programs
    - Runs on both ATLAS clusters + several other Tier3 sites
    - Focuses on batch processing of locally-distributed files
    - Runs since 2009 (<0.1% fault rate, used for most ATLAS publications done here)
  - **jCondor** (http://atlaswww.hep.anl.gov/asc/jcondor/)
    - Java applet for monitoring Condor submissions

See presentation by D. Malon on ATLAS-specific software development



ANL large-scale computer infrastructure

Argonne Leadership Computing Facilities (ALCF)

→ see next presentation

- **HEP leverages ALCF resources (like BlueGene/P)**
  - free of charge for ANL employers, but ALCF staff is a critical element
- **ASTRO group is involved in a simulation of the large scale structure of the Universe with the FLASH and MC^3 codes (level of ~0.4 FTE)**
  - BlueGene/P optimization
  - Scaling studies

see talk by K.Heitmann/S.Habib

Comment:

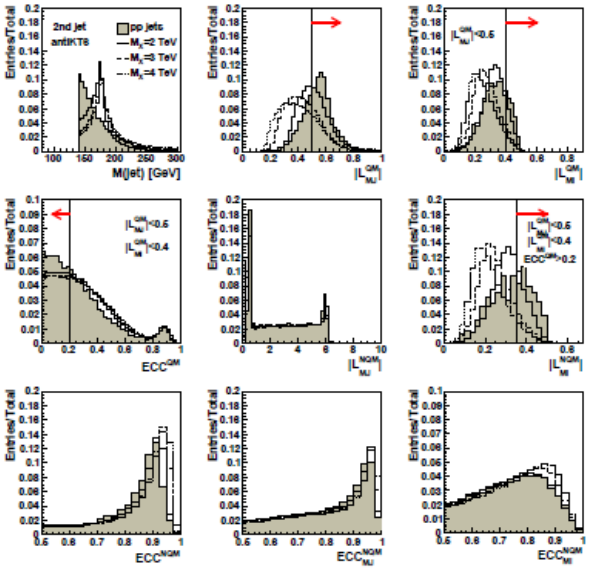
For many tasks (ATLAS, CDF, neutrino), BlueGene/P was difficult to use since it required a certain technical/intellectual investments to match HEP analysis-specific requirements:

- Low RAM per core (example: ATLAS requires >2 GB/core)
- No necessary libraries, not SL5/6 platform
- No optimal I/O for parallel analysis jobs
- HEP analysis jobs are usually rather short (~1 day/CPU) & require interactivity

Software solutions for general scientific HEP computation

- Note: we do not have any specific funds for such development
- “We made them, since we need them”

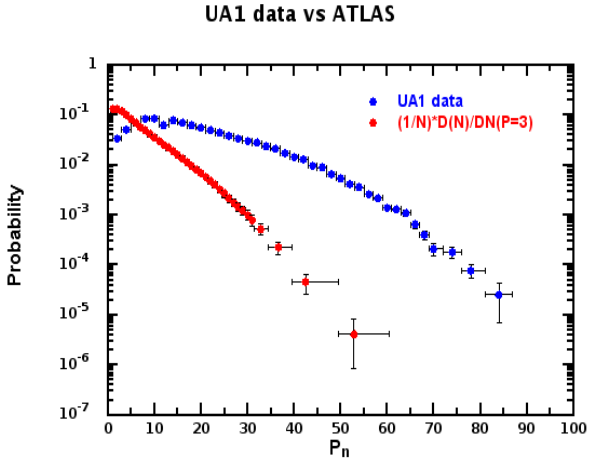
- **StatShapes** (http://atlaswww.hep.anl.gov/asc/statshape/)
  - C. Levy (summer student) + S.C.
  - a library for jet-shape reconstruction
  - ~20 jet-shape variables
  - 2 papers in PRL
  - moved to ATLAS code



S.C., C.Levy, J.Proudfoot, R.Yoshida, Phys.Rev.D82:094029,2010

Software solutions for general scientific HEP computation

- **RunMC** (http://projects.hepforge.org/runmc/) (S.C.)
  - C++/ROOT interface to legacy FORTRAM MC models (6 in total)
  - Contains Java GUI
- **jHepWork** (http://jwork.org/jhepwork/) - general-purpose data-analysis framework (S.C.)
  - Based on #1 programming language – Java
  - Jython (i.e. Python) used as as interface language to call Java libraries
  - Based on Jaida/FreeHEP libraries developed at SLAC
  - Rather popular outside HEP (~95% Windows), but limited popularity in HEP
  - According to TechSource review:
    - among 5 best open-source data mining software (ROOT did not make it)



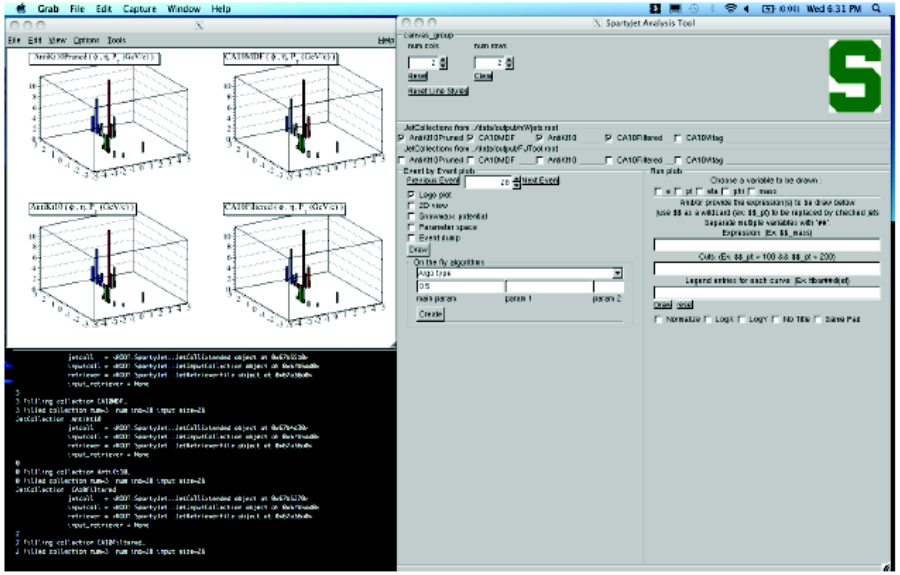
Used by **Durham HepData HEP database** for data export
Advantages: multiplatform + keeps track of statistical and systematical errors in a single data container

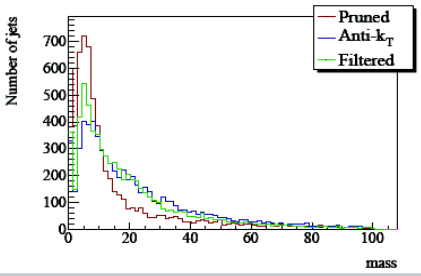
jHepWork output using Durham-generated Jython macro file

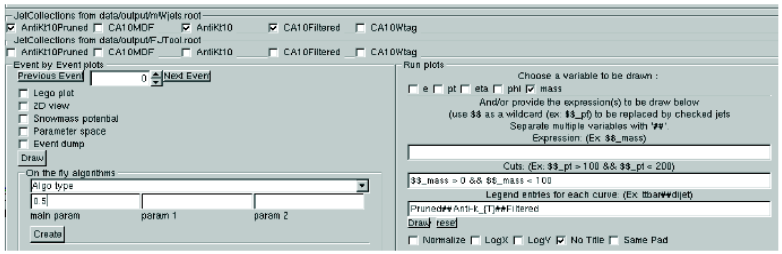
Software solutions for general scientific HEP computation

Software by ASC visitors:

- **SpartyJet** (http://projects.hepforge.org/spartyjet/)
  - B.Martin (ANL visitor from Mich. state) + others
  - Jet-algorithm library with GUI and many plugins
  - Contains all main jet algorithms







3-year plan. General infrastructure support

Our strategic plan depends on the funding model

All computing infrastructure support at ANL is from the research & R&D program

Main goal of HEP general computer support using the current budgetary model:

To continue to maintain the WINDOWS and LINUX clusters and the desktop environment

- No room for expansion using the current budgetary model. No extra FTE.
- Currently: we have ~$25k for hardware maintenance support for the HEP division
- This is not enough: For example:
  - the overall cost of only ATLAS computing infrastructure ~$250k
  - need at least $30k/year to maintain ATLAS clusters and to upgrade components on a 4-5 year replacement cycle
  - Dell maintenance will stop in 2 years

Current model is not optimal:

Uses research time to support general-purpose computing

3-year plan. A wish list: General infrastructure

- **Consolidate all hardware**
  - 3 LINUX clusters, Web servers, WINDOWS servers etc
- **Build (or reuse) a single computer room in Bld. 360 for 3 LINUX clusters**
  - professional cooling
  - uninterruptible power supplies
  - professional support (1FTE for LINUX general support)
- **20% of FTE's of 3 researches can be reduced to 5-10%**
  - shift their effort from “general computing” to “experiment-specific computing”
  - do not waste their unique capabilities as researches on general computing!
- **All of this will require:**
  - adequate hardware maintenance budget (~$100k/year for the division)
  - 1 FTE LINUX admin support
  - Funds to build a professional computer room

Summary

On the technical side:

1) we need to explore suitability of ANL general-purpose clusters
2) moving towards >24 core servers with SSD for interactive data processing for Tier3
   - smaller number of servers, less cooling/power consumption → less support

Next:

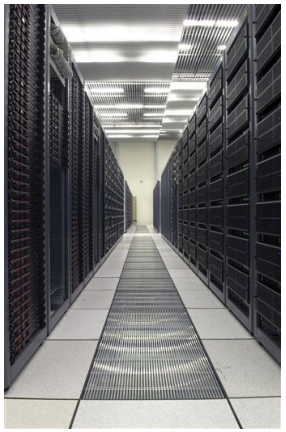
- ATLAS Computing — D.Malon
- Computational Cosmology — K.Heitmann/ S.Habib
- Overview of ANL computing — J.Bernstein

Backup

Software development

As one of DOE's flagship computing facilities, Argonne is poised for a leadership in DOE HEP's Cosmic Frontier

- We will base the effort on a cosmological simulation and analysis framework for large-scale structure probes of cosmology, aimed at upcoming surveys such as DES and LSST.
- **Simulation focus**:
  - developed new HACC (Hardware Accelerated Cosmology Code) framework, designed to scale up to the exascale
  - HACC has been chosen as an Early Science Project on Argonne's new 10 PFlops BG/Q system Mira, with an allocation of 150M CPU-hrs.
  - investigate melding of HACC and CCF code, and observational data archives into a single, integrated, interactive prediction
  - ALCF hired 2 employees to leverage this effort (S.Habib and K.Heitmann)



Collaboration with other laboratories

- **Very strong in all physics and R&D related issues**
- **Example:**
  - ***ATLAS analysis support center:***
    - *4 Joint Jamborees with BNL, 2 Joint Jamboree with LBNL*
    - *Coordination with BNL/LBNL for all issues related to USATLAS Tier3 support*
    - *Established collaboration with Esnet to solve networking issues*

- **Particularly strong collaboration by Argonne ATLAS software team with LBNL and BNL**
  - *with LBNL on ATLAS software architecture, I/O framework, performance optimization*
  - *with BNL on data management and event store persistence technologies, bytestream (raw data format) and AthenaPOOL (ROOT-based derived data technology)*